

# HIV Database Workshop

[www.hiv.lanl.gov](http://www.hiv.lanl.gov)

[seq-info@lanl.gov](mailto:seq-info@lanl.gov)

**Presenters:** Bette Korber, Will Fischer and Brian Foley  
**Database Pls:** Bette Korber, Thomas Leitner, Karina Yusim

**Additional database staff:** Werner Abfalterer,  
Peter Hraber, Elisabeth Sharon Fung, Robert Funkhouser, Jenni Macke,  
Kumkum Ganguly, James Szinger, and Hyejin Yoon

**Project Officer:** Stuart Shapiro, NIAID, NIH

*Theoretical Biology and Biophysics, T-6  
Los Alamos National Laboratory*



## Workshop Topics

**HIV Sequence Database and Immunology Database**  
Brian Foley, Karina Yusim

### **Session 2**

Thursday,  
March 13  
11:15 – 12:30

*Immunology database introduction*  
*Epitope maps and epitope summary tables*  
*T-cell epitope search*  
*T-cell epitope variants*  
*Antibody search*  
*List of most broadly neutralizing antibodies*  
*HIV/SIV sequence locator tool*  
*QuickAlign – Align an epitope to the database alignments*  
*CATNAP*  
*ELF – epitope location finder*  
*Peptgen – Design peptides for reagent development*

*Mosaic Vaccine Maker, Epicover, and Posicover*  
- generate candidate vaccines  
- estimate epitope coverage  
- determine regional epitope coverage

# Immunology Database Overview

- Incorporates published HIV T cell (CTL, T-helper) epitope and Antibody information (emphasis on monoclonals)
- Key information regarding what is learned about epitopes and mAbs in each paper is included
- Types of data recorded:
  - ☐ Epitope sequence and location: HXB2 numbering, subtype
  - ☐ Natural infection or vaccine
  - ☐ Host HLA or MHC
  - ☐ Ab isotype, binding region, species
  - ☐ Notes summarize main findings



HIV molecular immunology database

DatabasesSearchToolsProductsPublications

Search Site

HIV Molecular Immunology Database

The HIV Molecular Immunology Database is an annotated, searchable collection of HIV-1 cytotoxic and helper T-cell epitopes and antibody binding sites.

Search Interfaces

- [CTL/CD8+ search](#)
- [T Helper/CD4+ search](#)
- [Antibody search](#)
- [CTL variant search](#)
- [Search help](#)
- [Variant search help](#)

Database Products

- [All Database products and publications](#)
- [Epitope maps](#)
- [Epitope tables](#)
- [Epitope alignments](#)
- [CTL epitope variants and escape mutations](#)
- [Neutralizing antibody resources](#)
- [The HIV Molecular Immunology Compendium](#)
- [About the HIV Molecular Immunology Database](#)
- [How to cite this database](#)
- [Frequently-asked Questions \(FAQ\)](#)

Tools and Data Sets

- [Tools & Links](#) for immunologists
- [SIV Epitopes \(PDF\)](#) review article summarizing known SIV epitopes
- [Identifying HLA-Associated Polymorphisms in HIV-1 \(PDF\)](#) review article summarizing HIV polymorphism associated with escape mutations. Also a [table of polymorphisms](#).
- [HLA-TEM HLA Typing and Epitope Mapping Data Sets](#)
- [Standardized Assessments of Neutralizing Antibodies for HIV/AIDS Vaccine Development](#) Assay protocols from Duke Central Reference Laboratory

News

News Archive

CATNAP: Compile, Analyze, and Tally Neutralizing Antibody Panels

Our new tool, CATNAP, compiles IC<sub>50</sub> and IC<sub>80</sub> neutralization panel data for HIV-1 broadly neutralizing antibodies. It provides tools for meta-analysis of neutralization panel data and viral Envelope sequences. 26 February 2014

HIV Genome Browser

Our new visualization tool, Genome Browser, is a customization of JBrowse designed to provide graphic views of the HIV genome and proteome. It incorporates many sources of data from the HIV Sequence and Immunology Databases, including epitopes, entropy, functional domains, and many features of interest. 26 February 2014

Last modified: Wed Sep 29 09:44:51 MDT 2010

Questions or comments? Contact us at [immuno@lanl.gov](mailto:immuno@lanl.gov)

The logo for Los Alamos National Laboratory, featuring a stylized 'A' and the text 'Los Alamos NATIONAL LABORATORY'.

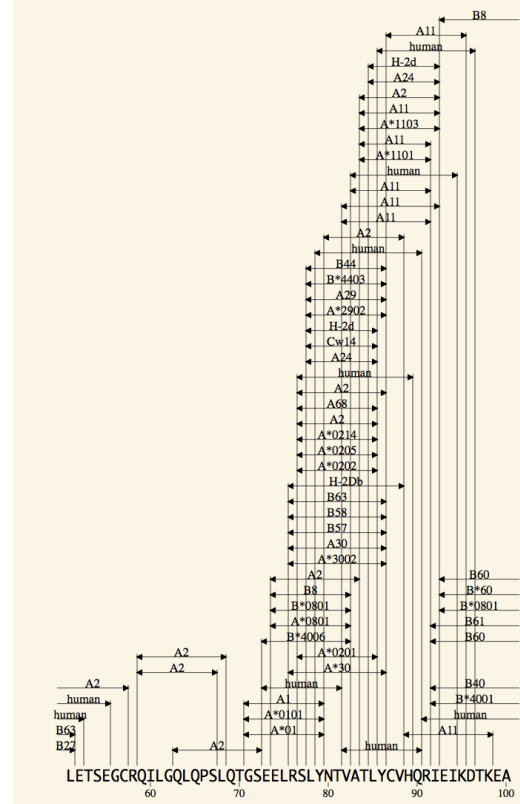
# Immunology Database Products

- Epitope maps (species/HLA for T cell epitopes; species/MAb name for Ab)
- Epitope summary tables:
  - All CTL and Helper epitopes and Ab binding sites
  - Variants of CTL epitopes
  - Christian Brander keeps an “A list” of HIV CD8+ T-cell epitopes – experimentally validated optimal epitopes with known HLA presenting molecules, will be updated soon
  - “B list” – a comprehensive list of all unique epitopes in the database (unknown HLA, boundaries not fully defined...)
  - All antibodies organized by protein and binding region
  - Antibody “A-list” – a table of the most broadly neutralizing MABs, with links to sequence and structure
- Tools for immunologists
- Yearly HIV Molecular Immunology Compendium



## p17 CTL/CD8+ Epitope Map

- Epitopes up to 14 aa long are mapped on HXB2
- HXB2 sequence may differ
- Epitopes with identical boundaries and HLA fields are included in the maps only once
- The epitope maps are interactive!



## CTL/CD8+ Epitope Summary (B-list)

- List of all epitopes up to 21 aa long
- Unlike epitope maps that show epitope locations, here each epitope sequence is shown

Epitope	Protein	HXB2 Location	Subtype	Species	HLA
MGARASVLSG	p17	1-10	CRF01_AE	human	
ASVLSGGEL	p17	5-13	B	human	
ASILRGGKLDK	p17	5-15	C	human	
SVLSGGQLDR	p17	6-15	B	human	A11
LSGGELDRWEK	p17	8-18		macaque	
GELDRWEKI	p17	11-19	B	human	B*4002, B40
GQLDRWEKI	p17	11-19	B	human	
GKLDSEWKIRLR	p17	11-22	A, CRF01_AE, CRF02_AG	human	
GKLDSEWKIRLR	p17	11-22	CRF01_AE	human	
ELDRWEKIRL	p17	12-21	B, C	human	B63
EKILRPGGKKYKL	p17	17-31		human	B27, B7
KILRPGGK	p17	18-26	A, A1, B, CRF01_AE	human, transgenic mouse	A*0301, A11, A3, B27, B7
KILRPGGKK	p17	18-27	B, C, multiple	human	A*0301, A11, A3, B27
KILRPGGKKYKL	p17	18-31		human	A3, B62

## Best-defined CTL/CD8+ Epitope Summary (A-list)

- Selective list of best defined epitopes as described by Christian Brander and colleagues

Epitope	Protein	HXB2 Location	Subtype	Species	HLA
GELDRWEKI	p17	11-19		human	B*4002
KILRPGGK	p17	18-26		human	A*0301
ILRPGGKK	p17	19-27	B	human	B*2705
RLRPGGKKK	p17	20-28		human	A*0301
RLRPGGKKKY	p17	20-29	B	human	A*0301
GGKKYKLLK	p17	24-32	B	human	B*0801
KYKLKHIVW	p17	28-36	B	human	A*2402
HLVWASREL	p17	33-41		human	Cw*0804
LVWASRELERF	p17	34-44		human	A30
WASRELERF	p17	36-44	B	human	B*3501
ELRSLYNTV	p17	74-82		human	B*0801
RSLYNTVATLY	p17	76-86	B	human	A*3002, B58, B63
SLYNTVATL	p17	77-85	B	human	A*0201, A*0202, A*0205
LYNTVATL	p17	78-85		human	Cw14
LYNTVATLY	p17	78-86		human	A*2902, B*4403
TLYCVHQK	p17	84-91		human	A*1101
IEIKDTKEAL	p17	92-101		human	B*4001
NSSKVSQNY	p17	124-132	B	human	B*3501



## Immunology Database: Search

- T Cells
  - ☐ Cytotoxic T Lymphocytes (CTL)
  - ☐ Helper T Lymphocytes (T-helper)
  - ☐ Organization is identical for CTL and T-helper
  - ☐ One reference per entry, epitope/HLA combinations are often repeated
- B Cells (Antibodies)
  - ☐ One entry for each monoclonal antibodies
  - ☐ Many references per entry (> 400 for some well studied MAbs)



# CTL/CD8+ T-cell Search

- Can search by HIV protein, Epitope Sequence, Subtype, Immunogen, Vaccine Details, Species, presenting MHC/HLA, Author, Country, Keywords
- Can now search on epitope location and find fuzzy matches, overlaps and embedded epitopes
- Search example:
  - SLYNTVATL – 254 entries
  - To narrow the search use keyword “escape” – 32 entries
- Additional information provided in the entry:
  - Location, Donor MHC/HLA, experimental methods, Notes
  - CTL epitope variants if studied in the paper
  - Link to all entries for a reference
  - PubMed links to papers
  - Link to Epitope Maps
  - Link to Epitope Alignment (Extracted from HIV-sequence database, includes subtype, country and year of sampling)



## CTL/CD8+ T-cell Search

Search for

Epitope: ISPRTLNAW

First Author: Pillay

### Search CTL/CD8+ T-Cell Epitope Database

<a href="#">HIV protein</a>	Proteins with <a href="#">defined epitopes</a> - ALL - p17 p17-p24 p24 p24-p2p7p1p6	Proteins with <a href="#">undefined epitopes</a> - ALL - Gag Gag/Pol Pol Vif
<a href="#">HXB2 location</a>	<input type="text"/> - <input type="text"/>	
<a href="#">Epitope</a>	ISPRTLNAW	
<a href="#">Epitope name</a>	<input type="text"/>	
<a href="#">Record number</a>	<input type="text"/>	
<a href="#">Subtype</a>	- ALL -	
<a href="#">Immunogen</a>	- ALL - computer prediction HIV-1 and GBV-C co-infection HIV-1 and HCV co-infection HIV-1 exposed seronegative HIV-1 infected monocyte-derived HIV-1 infection	
<a href="#">Vaccine details</a>	Vaccine type: - ALL - Vaccine strain: - ALL - if Immunogen is Vaccine Vaccine component: - ALL - Adjuvant: - ALL -	
<a href="#">Species</a>	- ALL -	
<a href="#">MHC/HLA</a>	- ALL - A*01 A*0101 A*02 A*0201 A*02.01 A*020101	
<a href="#">Author</a>	Pillay	
<a href="#">Country</a>	- ALL -	
<a href="#">Keywords</a>	- ALL - acute/early infection adjuvant comparison antagonism antibody binding site definition and exposure assay development, comparison, standardization, improvement autologous responses	
<a href="#">Note</a>	<input type="text"/>	

[Click for Search Help](#)

## Search CTL/CD8+ T-Cell Epitope Database

Found 1 matching record:

### Displaying record number 53832

<a href="#">HXB2 Location</a>	p24(15-23)
<a href="#">Author Location</a>	Gag(147-155)
<a href="#">Epitope</a>	ISPRTLNAW
<a href="#">Subtype</a>	C
<a href="#">Species (MHC/HLA)</a>	human(B57)
<a href="#">Immunogen</a>	HIV-1 infection
<a href="#">Donor MHC/HLA</a>	A*3001, A*66, B*4201, B*5802, Cw*0602, Cw*1701; A*66, A*68, B*57, B*5802, Cw*0602, Cw*0701
<a href="#">Country</a>	South Africa
<a href="#">Experimental methods</a>	CD8 T-cell Elispot - IFN $\gamma$
<a href="#">Keywords</a>	epitope processing, responses in children, mother-to-infant transmission, escape, acute/early infection

[p24 Epitope Map](#)

[Epitope Alignment](#)  
[Show epitope variants](#)

Variant details with  
annotator's notes

### Notes

- HIV-specific CTLs in infants were shown to be able to select for viral escape variants early in life, despite a lack of escape with the same CTL specificity in the mother. Infant CTL responses may be compromised by transmission of escape variants that arose in the mother and also those that arose in the father, if the father was the source of the mother's infection.
- ISPRTLNAW is the C consensus form of the epitope and was the autologous form in the mother, and was transmitted to her infant. By 33 weeks a new dominant form of the epitope had emerged in the infant, mSPRTLNAW, and two additional variants had arisen, one with a substitution proximal to the epitope, pISPRTLNAW, and ISPRTLNAW.

### References

Pillay2005 Thillagavathie Pillay, Hua-Tang Zhang, Jan W. Drijfhout, Nicola Robinson, Helen Brown, Munira Khan, Jagadesa Moodley, Miriam Adhikari, Katja Pfafferoth, Margaret E. Feeney, Anne St. John, Edward C. Holmes, Hoosen M. Coovadia, Paul Klennerman, Philip J. R. Goulder, and Rodney E. Phillips. Unique Acquisition of Cytotoxic T-Lymphocyte Escape Mutants in Infant Human Immunodeficiency



### Displaying record number 53832

## Variants details

<a href="#">HXB2 Location</a>	p24(15-23)	<a href="#">p24 Epitope Map</a>
<a href="#">Epitope</a>	ISPRTLNAW	<a href="#">Epitope Alignment</a>
<a href="#">Variants</a>	mSPRTLNAW ISPRTLNAW pISPRTLNAW	escape documented in this paper diminished response not determined
<a href="#">Species (MHC/HLA)</a>	human(B57)	

Can go back to epitope entry

#### Variant Details

Showing all 3 variants.

Variant ID.	1413
Epitope Seq.	ISPRTLNAW
Variant Seq.	mSPRTLNAW
Mutations	I/M
Epitope Location	I1M
HXB2 Location	I15M
Mutation Type	E: escape documented in this paper
Method	CD8 T-cell Elispot - IFN $\gamma$ , Sequence
Note	This is de novo variant seen in infant by week 33 of age. The index peptide was recognized, but not the variant.

Mutation type

Note describing  
why the variant  
was designated  
particular  
mutation type

Variant ID.	1414
Epitope Seq.	ISPRTLNAW
Variant Seq.	ISPRTLNAW
Mutations	I/L
Epitope Location	I1L
HXB2 Location	I15L
Mutation Type	DR: diminished response
Method	CD8 T-cell Elispot - IFN $\gamma$ , Sequence



# Summary table of ~ 2800 epitope variants

CTL/CD8+ Epitope Variant Details															
Download CTL/CD8+ epitope variant details as CSV or XLS files.															
<a href="#">Link of Mutation types</a>															
Data last updated at 2013-01-25 11:58:16:07															
Epitope ID	Epitope Name	Variant ID	Subtype	Epitope Subtype	Variant Subtype	Protein	HXB2 start	HXB2 end	HLA	Epitope	Variant Epitope	Mutation (epitope)	Mutation (protein)	Mutation Type	Note
54532	A114	1016	B	B	A, M-group	p17	5	19		ASVLSGGELDRWEKI	ASVLSGGALDAWEKI	R11A, E8K	R15A, E12K	SNFS	No cross-recognition of this variant was seen across clades or intra-clade central sequences.
54532	A114	1017	B	B	C	p17	5	19		ASVLSGGELDRWEKI	ASLVGGALDAWEKI	R11K, V3I, S5R, E8K	R15K, V7L, S5R, E12K	SNFS	No cross-recognition of this variant was seen across clades or intra-clade central sequences.
54532	A114	1018	B	B	B	p17	5	19		ASVLSGGELDRWEKI	ASVLSGGALDAWEKI	R11K, E8K	R15K, E12K	SNFS	No cross-recognition of this variant was seen across clades or intra-clade central sequences.
54532	A114	1019	B	B	B	p17	5	19		ASVLSGGELDRWEKI	ASVLSGGELDAWEKI	R11K	R15K	SNFS	No cross-recognition of this variant was seen across clades or intra-clade central sequences.
33591	Gag 1.2	34	B		CRF02_AG	p17	8	18		LSGGELDRWEK	LSGGALDAWEK	E5K, R8A	E12K, R15A	SNFS	Intracellular cytokine staining, T-cell Elispot. CRF02 form, LSGGALDAWEK, does not cross-react with the B clade LSGGELDRWEK elicited response.
33844	GP	1569	B			p17	11	19	B40	GELDRWEKI	GELDRWAKI	E7K	E17K	DR, LE	This variant from the HXB2 sequence was present in the restricting HLA-B40-carrying mother, M-1002, but was never detected in her non-HLA-B40-carrying infant, P-1031. Decreased recognition of the E7K variant relative to the index epitope was seen in the mother.
56027	GR(p17)	1903	B	B	B	p17	11	19		GQLDRWEKI	GalDRWEKI	Q2E	Q12E	ND	This Asian B Clade optimal epitope differs from the consensus B at one position. It is predicted to be HLA-B40 restricted. Experimentally, B clade consensus peptide was used to challenge CTL response in subjects commonly carrying the Asian B-type epitope.
55632		11		A, CRF02_AG, A, AG, CRF01_AE	AE	p17	11	22		GKLSDWEKRLR	GKLDWEKRLR	S5A	S15A	SSF	1 subject responded to peptide GKLDWEKRLR from subtypes CRF02_AG and A and to peptide GKLSDWEKRLR from subtype CRF01_AE.
54629	GAG-03	1957	B	B	C	p17	17	34		EKILRPGGKKYKYL	EKILRPGGKKYLMKL	K12H, K28H, R14H	K28H, R30H	SSF	This Clade C consensus synthetic peptide variant from an immunodominant region, differs from the immunodominant Clade B consensus at 2 amino acids (11, 15) and both were recognized by subtype B-infected subjects.
33601	KK9	31	B			p17	18	26	A3	KILRLPGGK	KILRLPGGq	K9Q	K26Q	E, P	Variant inhibits processing, resulting in rapid decline in the K99 specific CD8+ T-cell response.
32770	KK9	153	B			p17	18	26	A3	KILRLPGGK	KILRLPGGr	K9R	K26R	SF	Flow cytometric T-cell cytokine assay. KILRLPGGK was recognized by 3 patients. The autologous sequence in one patient was KILRLPGGq which induced high frequency response.
35233		790	B, CRF01_AE		B	p17	18	26	A3	KILRLPGGK	KILRLPGGr	K9R	K26R	IE	Flow cytometric T-cell cytokine assay. Patient was superinfected with three strains, B1, B2 and CRF01_AE. This variant developed in B1 to include 42% of the viruses within 4 years.
This variant was used in DePrent and															



## Antibody Search

- Can search by
  - HIV protein, Epitope Sequence, Subtype, Immunogen, Vaccine Details, Species, presenting MHC/HLA, Author, Country, Keywords
  - MAb ID (Ab lists by name and by binding type are provided)
  - Ab type (by binding site, for example binding to similar region like V3 or near a common functional domain like CD4 binding site CD4Bs)
  - Isotype
- Search examples:
  - 2F5 – 1 record with 463 references
  - Ab type: gp120 CD4BS – 200 records



## Search Antibody Database

# Antibody Search

<a href="#">HIV protein</a>	<a href="#">Proteins with defined epitopes</a> - ALL - p17 p17-p24 p24 p24-p2p7p1p6	<a href="#">Proteins with undefined epitopes</a> - ALL - p24 Gag RT Pol
<a href="#">HXB2 location</a>	Results overlap with query location	
<a href="#">Epitope</a>	Results contain query sequence	
<a href="#">Record number</a>		
<a href="#">MAb ID</a>	(List by name) (List by type)	
<a href="#">Subtype</a>	- ALL -	
<a href="#">Immunogen</a>	- ALL - anti-idiotype autoimmune disease HIV-1 exposed seronegative HIV-1 infection HIV-2 infection in vitro stimulation or selection	
<a href="#">Vaccine details if Immunogen is Vaccine</a>	<a href="#">Vaccine type</a> - ALL - <a href="#">Vaccine strain</a> - ALL - <a href="#">Vaccine component</a> - ALL - <a href="#">Adjuvant</a> - ALL -	
<a href="#">Ab Type</a>	- ALL - C-domain C-HR C-term Env oligomer flap region gp120 adjacent to CD4BS	
<a href="#">Species</a>	- ALL -	
<a href="#">Isotype</a>	- ALL - IgA IgA1 IgA2 IgA2a IgE IgG	
<a href="#">Author</a>	Search only for <input type="radio"/> First <input type="radio"/> Last author <input checked="" type="radio"/> Show only this author's references <input type="radio"/> Show all references	
<a href="#">Country</a>	- ALL -	
<a href="#">Keywords</a>	<input checked="" type="radio"/> Show only notes containing selected keyword(s) <input type="radio"/> Show all notes	
<a href="#">Note</a>	<input checked="" type="radio"/> Show only notes matching this text <input type="radio"/> Show all notes	
<input type="button" value="Search"/> <input type="button" value="Reset"/> <a href="#">Click for Search Help</a>		

Can search by HXB2 location, Find overlaps, fuzzy matches Embedded epitopes

Can show only notes and references containing selected keywords or user's text (as apposed to showing matching Ab entries with all notes)



# Antibody Search

Found 1 matching record:

## Displaying record number 815

[MAb ID](#) 2F5 (IAM 2F5, IAM-41-2F5, IAM2F5, c2F5)  
[HXB2 Location](#) gp160(662-667)  
[Author Location](#) gp41(662-667 BH10)  
[Research Contact](#) Hermann Katinger, Institute of Applied Microbiology, Vienna, or Polymun Scientific Inc., Vienna, Austria  
[Epitope](#) ELDKWA  
[Ab Type](#) gp41 adjacent to cluster II, C-term, gp41 MPER (membrane proximal external region)  
[Neutralizing](#) L P  
[Species \(Isotype\)](#) human(IgG3κ)  
[Immunogen](#) HIV-1 infection

[gp160 Epitope Map](#)

[Epitope Alignment](#)

## Keywords

acute/early infection, adjuvant comparison, anti-idiotype, antibody binding site definition and exposure, antibody generation, antibody interactions, antibody sequence variable domain, assay development, standardization and improvement, autoantibody or autoimmunity, autologous responses, binding affinity, brain/CSF, co-receptor, complement, dendritic cells, drug resistance, enhancing activity, escape, genital and mucosal immunity, HAART, ART, HIV exposed persistently seronegative (HEPS), immunoprophylaxis, immunotherapy, immunotoxin, isotype switch, kinetics, mimics, mimotopes, mother-to-infant transmission, neutralization, rate of progression, responses in children, review, SIV, structure, subtype comparisons, supervised treatment interruptions (STI), therapeutic vaccine, vaccine antigen design, vaccine-induced immune responses, variant cross-recognition or cross-neutralization, viral fitness and reversion

## Notes

- 2F5: 2F5 neutralized infection of PBLs with various HIV-1 strains with high potency. However, 2F5 did not inhibit transcytosis of cell-free or cell-associated virus across a monolayer of epithelial cells. A mixture of 13 MAb directed to well-defined epitopes of the HIV-1 envelope, including 2F5, did not inhibit HIV-1 transcytosis, indicating that envelope epitopes involved in neutralization are not involved in mediating HIV-1 transcytosis. When the mixture of 13 MABs and HIV-1 was incubated with polyclonal anti-human γ chain, the transcytosis was partially inhibited, indicating that agglutination of viral particles at the apical surface of cells may be critical for HIV transcytosis inhibition by HIV-specific Abs. [Chomont2008 \(neutralization\)](#)
- 2F5: The lipid binding properties of 2F5, and the similarity to binding properties of anti-lipid mAbs, are discussed. Potential role of liposomes containing lipid A for induction of NABs to lipids of HIV-1 is reviewed. [Alving2008 \(autoantibody or autoimmunity, review\)](#)
- 2F5: A reference panel of recently transmitted Tier 2 HIV-1 subtype B envelope viruses was developed representing a broad spectrum of genetic diversity and neutralization sensitivity. The panel includes viruses derived from male-to-male, female-to-male, and male-to-female sexual transmissions, and CCR5 as well as CXCR4 using viruses. The envelopes displayed varying degrees of neutralization sensitivity to 2F5, with 14 of 19 envelopes sensitive to neutralization by this Ab. [Schweighardt2007 \(assay development, standardization and improvement, neutralization\)](#)
- 2F5: This review summarizes data on possible vaccine targets for elicitation of neutralizing Abs and discusses whether it is more practical to design





# Antibody “A-list”

- List of most broadly neutralizing antibodies – currently 45 MAbs (*work in progress*)
  - Links to papers, Ab sequences and structures
  - Notes on breadth of neutralization
  - Notes on Ab contact residues
  - Notes on heavy and light chain composition
- Under “Database products”
- [http://www.hiv.lanl.gov/content/immunology/tables/ab\\_best\\_neutralizing\\_summary.html](http://www.hiv.lanl.gov/content/immunology/tables/ab_best_neutralizing_summary.html)



Databases

Search

Tools

Products

Publications

Search Site


Summary of Best Neutralizing Antibodies

Download summary of best neutralizing antibodies as [CSV](#) or [XLS](#) files.

This is a list of most broadly neutralizing HIV antibodies, with links to papers, Ab sequences and structure, notes on breadth of neutralization, Ab contact or key residues and heavy and light chain composition. Note: this is a work in progress, so not all relevant papers and antibodies are listed.

Mab	Binding site	Author Journal Pmid	First paper	Breadth of neutralization with IC50<50 µg/ml	Breadth of neutralization with IC80 or IC90<50 µg/ml	Structure, PDB ID	Ab sequence	Heavy chain	Light chain	Germline Ab sequence	Ab binding affinity	Listings of antibody contact or key residues	
VRC01	CD4bs	<a href="#">Wu2010</a> Science 20616233	YES	91% of 190 isolates, representing major HIV-1 clades	86% of 190 isolates, representing major HIV-1 clades, with IC80		<a href="#">GI:294875838</a> -- heavy chain variable region  <a href="#">GI:294875848</a> -- light chain variable region	V: IGHV1-02*02 D: IGH D3-16*01 (or *02) J: IGHJ1*01 or IGHJ2*01	V: IGKV3-11*01 J: IGKJ2*01	Fig. S5	Bound strongly to RSC3 and gp120 and weakly to ΔRSC3, Fig. 2 and S4.		
		<a href="#">Zhou2010</a> Science 20616231					<a href="#">3NGB</a>				Fig. S12	Figs. 5, 6, S3	Env, defined by crystal structure: Fig S1. Antibody, defined by crystal structure: Fig. S9
		<a href="#">Wu2011</a> Science 21835983									Sequence, Figs. 1, S14, S18. Phylogenetic analysis, Fig. 5, Fig. S13		Antibody, defined by crystal structure, compared to key residues of other CD4bs antibodies, Fig. S4.
		<a href="#">Scheid2011</a> Science 21764753		100% of 118 isolates representing major HIV-1 clades								Fig. 3, Table S9.	Antibody, defined by crystal structure in Zhou2010, Fig. 4, Fig. S3, and Fig. S4 provide comparisons with other CD4bs Nabs.
		<a href="#">Walker2011</a> Nature 21849977		93% of 162 isolates representing major HIV clades	89% of 162 isolates representing major HIV clades, with IC90								

[http://www.hiv.lanl.gov/content/immunology/tables/ab\\_best\\_neutralizing\\_summary.html](http://www.hiv.lanl.gov/content/immunology/tables/ab_best_neutralizing_summary.html)  
 Epitope Summary Tables  
 Best Neutralizing Antibodies



# Tools for Immunologists

- **Sequence Locator** Finds epitope location on the reference genome
- **QuickAlign** Aligns amino acids or nucleotides against our alignments
- **PeptGen** Generates overlapping peptides for any protein
- **CATNAP**: Compile, Analyze and Tally NAb Panels
- **ELF** Epitope Location Finder
- **N-Glycosite** Finds N-linked glycosylation sites
- **Mosaic** Generates candidate vaccine protein cocktails
- **Heatmaps** Displays and organizes neutralization or other quantitative data.
- And more ...



<http://www.hiv.lanl.gov/content/immunology/tools-links.html> HIV molecular immunology database

Databases Search Tools Products Publications Search Site

## HIV Molecular Immunology Database: Tools & Links

### Tools Produced by the Los Alamos HIV Databases

- [QuickAlign](#) Align amino acids or nucleotides against our alignments
- [PeptGen](#) Generate overlapping peptides for any protein
- [PeptMap](#) Generate peptide maps in Fasta, HTML and PDF formats
- [Hepitope](#) Search for hopeful epitopes based on HLA enrichment
- [HLA Frequency Analysis Tools](#) Calculate HLA frequencies or HLA linkage disequilibrium in a population
- [ELF](#) Epitope location finder
- [Motif Scan](#) Scan alignments for HLA binding motifs
  - [HLA genotype/serotype dictionary](#)
  - [HLA genotype/motif dictionary](#)
  - [HLA supertype dictionaries](#)
- [HIV/SIV Sequence Locator Tool](#) Formerly the *HXB2 Numbering Engine*
- [SeqPublish](#) Produce pretty alignments for publication
- [BLAST](#) Search sequences against our annotated HIV sequences
- [Heatmap](#) Display a table of numbers using colors to represent the numerical values
- [Mosaic Vaccine Tool Suite](#) Design and assess polyvalent protein sequences for T-cell vaccines
- [N-Glycosite](#) Find N-linked glycosylation sites
- [Highlighter](#) Highlight matches and mismatches in a set of aligned sequences
- [HIV Genome Browser](#) Display HIV genome and proteome
- [CATNAP: Compile, Analyze and Tally NAb Panels](#)
- [All Tools](#) List of all software and tools in both the HIV sequence and Immunology databases

### External Tools for Epitope Prediction

- [BIMAS HLA Peptide Binding Predictions](#) Ranks potential n-mer peptides based on a predicted half-time of dissociation to HLA class I molecules
- [SYFPEITHI Epitope Prediction](#) Predicts the binding of a given amino acid sequence to a defined HLA type
- [PAProC](#) Predicts cleavages by human and yeast 20S proteasomes
- [PREDEP](#) MHC class I epitope prediction
- [MHCpred](#) Predicts MHC/peptide or TAP/peptide IC<sub>50</sub> binding values
- [Microsoft Research Epitope Predictor](#) Computes the probability that a given n-mer is a T-cell epitope restricted to a given HLA allele
- [Max Planck Institute](#)
  - [FRAGPredict](#) Prediction of proteasome cleavage
  - [MAPPP](#) Prediction of MHC-I antigenic peptide processing
- [Immune Epitope Database \(IEDB\)](#)
  - [MHC Class I Binding](#)
  - [MHC Class II Binding](#)
  - [NetMHC and NetMHCII](#) Predictors of antigen presentation and T-cell epitopes



# HIV/SIV Sequence Locator Tool

- Instantly computes position numbers of DNA or protein fragments relative to a reference strain (HXB2r for HIV-1, SMM239 for SIV)
  - Such numbers, often included in the literature, are frequently incorrect
- Shows the location of the sequence on an HIV map
- Presents protein translations of DNA sequences
- Can be used for input into the search interface, to align a new sequence you have generated with the database set
- Can also retrieve reference sequences
  - by coordinates (range of base or amino-acid positions)
  - by single position (retrieves flanking sequences)



<http://www.hiv.lanl.gov/content/sequence/LOCATE/locate.html>

## HIV Sequence Locator Tool

**Purpose:** This tool has several purposes. It can find the start and end coordinates (relative to the reference strain HXB2) of your input sequence(s) and show which genes or proteins it covers, along with a graphical view of the location of your sequence(s) relative to the reference sequence. The tool will display both the nucleotide sequence and protein translation of your input as it aligns to HXB2. It will also check the reverse complement of your input sequence, and report the orientation with the best match. Another use is to retrieve a section of the HXB2 reference sequence based on its coordinates.

**How to use:** To find the coordinates for your sequence, either upload or paste your sequence (any format) in the box below, or (for database sequences only) enter GenBank accession numbers. To retrieve the HXB2 sequence for a set of coordinates (see [HIV coordinate map](#)), enter the coordinates and choose the region. To retrieve the entire gene or protein, enter coordinate values of "1" and "end". To retrieve a single nucleotide or range with its surrounding 42-nucleotide sequence, enter the single coordinate in the "from" field and check the box. For more details, see [Sequence Locator Explanation](#).

### Useful Links:

[HXB2 numbering](#) | [SIVmm239 numbering](#) (review articles)  
[HXB2 spreadsheet](#) | [SIVmm239 spreadsheet](#) (spreadsheets with base-by-base annotation)

### Find the location of a sequence

Sequence type: ☒ Let program decide ☐ HIV ☐ SIV

Paste your input here  
[Sample Input]

SLYNTVATL

or upload your file

Paste or type a DNA or protein sequence here.

-- OR --

### Retrieve a region by its coordinates

Enter coordinates: from  to  (Enter '1' and 'end' to retrieve the entire region.)

Region:

Retrieve: ☒ Nucleotide or ☐ protein output

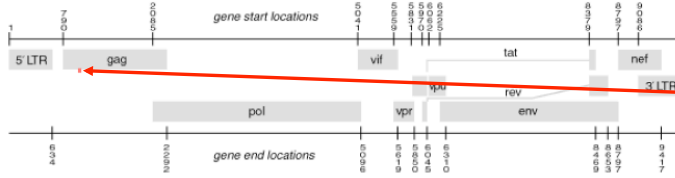
☐ include surrounding region

OR enter numeric coordinates here.



## Sequence Locator: “find my sequence”

Result for SLYNTVATL Query sequence



Location in genome mapped in red.

Table of protein regions touched by query sequence. AA = amino acid, NA = nucleic acid.

CDS	AA position relative to protein start in HXB2	AA position relative to query sequence start	AA position relative to polyprotein start in HXB2	NA position relative to CDS start in HXB2	NA position relative to HXB2 genome start
Gag	77 → 85	1 → 9	NA	229 → 255	1018 → 1044
p17	77 → 85	1 → 9	NA	229 → 255	1018 → 1044

Numeric coordinates useful for entry on search form

Alignment of the query sequence to HXB2 (Similarity 100.0%):

Query SLYNTVATL 9  
 :::::~  
 HXB2 SLYNTVATL

DNA and protein sequence displayed



## Sequence Locator: “Retrieve from coordinates”

-- OR --

Retrieve a region by its coordinates

Enter coordinates: from 77 to 85 (Enter '1' and 'end' to retrieve the entire region.)

Region: Gag

Retrieve: ☐ Nucleotide or ☒ protein output

☐ include surrounding region

Submit Reset

Include surrounding region

Reference Strain	Type	Region	Start	End
HXB2	pro	complete	77	85
Retrieved Sequence:				
SLYNTVATL				

Reference Strain	Type	Region	Start	End
HXB2	pro	complete	56	106
Retrieved Sequence:				
GCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEE				

50 aa long stretch



# QuickAlign

- Generates an alignment of your HIV-1 amino acid or nucleotide sequence against our web alignments
- Can be used to align epitopes, functional domains, or any protein or nucleotide region of interest
- Calculates frequency of variants to the query sequence and summarizes both by subtype and all subtypes together
- Calculates frequency of amino acid or nucleotide by position and summarizes both by subtype and all subtypes together



## QuickAlign

formerly EpiAlign and Primalign

Purpose: Align a desired region from our [Web alignments](#), with or without user-provided sequence(s). Details below.

### Retrieve alignment(s) based on sequence

Paste your sequence(s) here  
[\[Sample Input\]](#)

or upload sequence file

-- OR leave both fields above blank, and --

### Retrieve alignment(s) based on coordinates

Sequence coordinates  start  end

Gene/region/protein

### Options

Organism ☒ HIV1 ☐ HIV2 ☐ SIV

Sequence type ☐ nucleotide ☐ protein ☒ let program decide

[Alignment type](#) to use

[Delete Gaps](#) and shift sequence toward C-terminus (protein only) ☐ yes ☒ no

Display [wide output](#) ☐ yes ☒ no

Calculate [frequency by position](#) ☒   %

Include [surrounding region](#) ☐

The logo for Los Alamos National Laboratory, featuring a stylized 'A' and the text 'Los Alamos NATIONAL LABORATORY'.

## QuickAlign: example of output

- Query peptide:  
SLYNTVATL
- Sequence names include subtype, country and year of sampling
- Identical sequences are shown in red

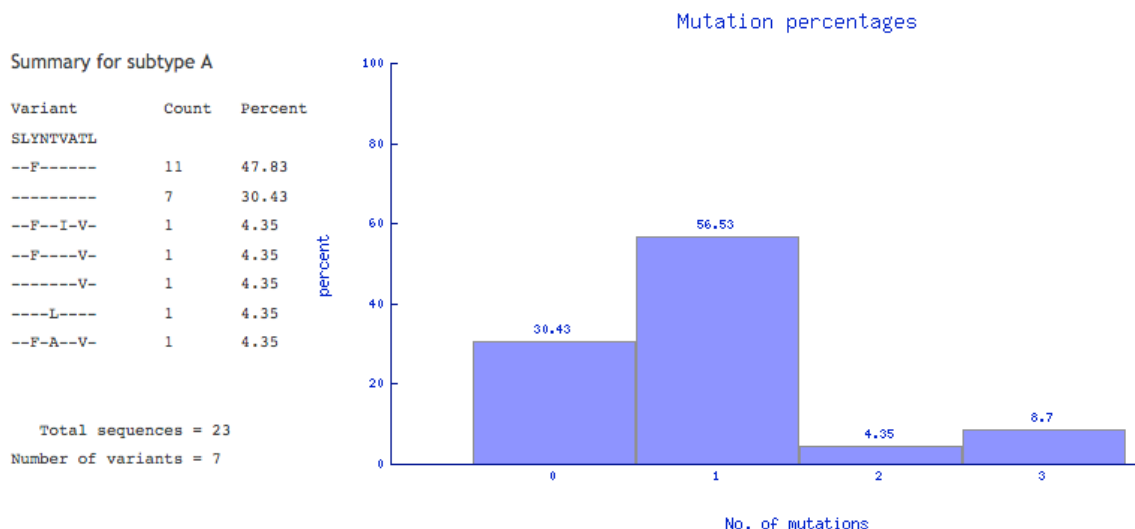
Query:	SLYNTVATL
Query Length:	9
HXB2 Location:	Gag 77-85 = p17 77-85
Alignment:	GAG, 458 sequences

Summarize

Query	SLYNTVATL
A1.KE.86.ML170	--F-----
A1.KE.94.Q23	--F-----
A1.SE.94.SE7253	--F----V-
A1.SE.94.SE7535	-----
A1.SE.95.SE8538	-----
A1.SE.95.SE8891	-----
A1.SE.95.UG8E8131	-----
A1.TZ.97.97T203	--F----V-



## QuickAlign: sequence variant summary



- Variant frequency summary by subtype and all subtype together



# QuickAlign: Frequency by position

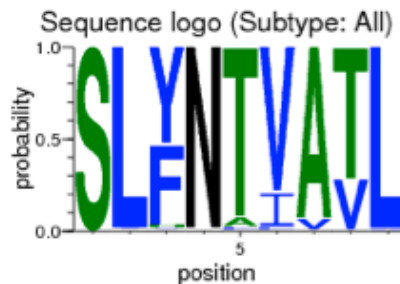
Frequency by position

[Go to top](#)

[See full raw counts](#)

cutoff: 95%

Position	Percentage and raw count of non-gap	Non-gap/total (percentage)
1	S: 99.90% (3113) other: 0.10% (3)	3116/3119 (100.00%)
2	L: 98.90% (3068) other: 1.10% (34)	3102/3119 (99.55%)
3	Y: 52.71% (1633) F: 43.77% (1356) other: 3.52% (109)	3098/3119 (99.42%)
4	N: 99.68% (3104) other: 0.32% (10)	3114/3119 (99.94%)
5	T: 92.86% (2887) A: 5.05% (157) other: 2.09% (65)	3109/3119 (99.78%)
6	V: 79.35% (2448) I: 18.15% (560) other: 2.50% (77)	3085/3119 (99.01%)
7	A: 92.95% (2889) V: 6.53% (203) other: 0.51% (16)	3108/3119 (99.74%)
8	T: 72.52% (2254) V: 27.06% (841) other: 0.42% (13)	3108/3119 (99.74%)
9	L: 99.00% (3078) other: 1.00% (31)	3109/3119 (99.78%)



## Neutralizing Abs: state of the field

- Despite 30 years of research, there is no HIV-1 vaccine, but recent developments offer a new hope for a protective immunization with neutralizing capacity
- In addition to the overall genetic variability and recombination events, HIV-1 Env spike escapes neutralizing antibody response through indels, hyper-variable loops, extensive glycosylation and conformational masking of vulnerable epitopes
- Only 4 (and not very potent or broad) cross-reactive neutralizing MAbs were known for more than 20 years of HIV-1 research
- During last 5 years several dozens of potent and broad NAbs were isolated, based on
  - New highly accurate neutralization assays and panels of 100s of diverse HIV isolates
  - New techniques to screen sera from many HIV-infected individuals to find elite neutralizers and clone Abs from their B cells
  - Novel Ab selection, screening and isolation approaches, including PCR amplification from single B-cells, structure-guided Env bait design, new PCR primers to target more conserved regions of immunoglobulin genes, next generation sequencing etc.
- The results of large neutralization panels can allow powerful meta-analyses to find antibody neutralization signatures and sites of vulnerability
- Most studies do not supply sufficient HIV sequence information: large Env panels are published without accession numbers, with a huge discrepancy in sequence names, making subsequent signature analysis of even one study difficult, let alone multiple studies
- The neutralization panels published as PDF tables, difficult to use



## Neutralizing Antibody Resources

### Tools

- [CATNAP: Compile, Analyze and Tally NAb Panels](#)  
Meta-analysis of neutralization panels for HIV-1 neutralizing antibodies.
- [HIV Genome Browser](#)  
A customization of jBrowse displaying genome and proteome features of HIV, including epitopes and neutralizing antibody features.

### Search interface

- [Neutralizing antibody contexts and features](#)  
Search for locations of important neutralizing antibody binding sites and other HIV-1 Env features.

### Tables

- [Neutralizing antibody contexts and features \(.xls\)](#)  
A summary of the information from the search interface above, presented in a single .xls spreadsheet. Each row of the table corresponds to one residue of HIV-1 Env, and each column represents a protein feature or set of known binding residues of broadly neutralizing antibodies. Loops and other features of Env are shown in the first 3 columns on the left. The entropy (sequence variability) of each residue is presented numerically and color coded. Abbreviated references are listed under each column heading, and full references are on the second page of the Excel file.
- [Best neutralizing antibodies](#)  
A table presenting the most broadly-neutralizing HIV-1 antibodies, with links to papers, Ab sequences, structures, notes on breadth of neutralization, locations of Ab contacts or key residues, and heavy and light chain composition.

Last modified: Wed Sep 29 09:44:51 MDT 2010

Questions or comments? Contact us at [immuno@lanl.gov](mailto:immuno@lanl.gov)

5  
Y

# CATNAP

## Compile, Analyze and Tally NAb Panels

CATNAP:  
Theoretical approximation

*By Peter Hraber*





### CATNAP

Compile, Analyze and Tally NAB Panels

**Purpose:** To provide easy access to data associated with HIV-1 neutralizing antibodies, including neutralization panel data, sequences, and structures. See [Explanation](#).

Select by antibody and virus
Select by study

**Antibodies**

# of Abs = 82

Select	Name	Donor	# of viruses tested
<input type="checkbox"/>	10-1074	Donor 17	119
<input type="checkbox"/>	10-996	Donor 17	119
<input type="checkbox"/>	10E8	Donor N152	181
<input type="checkbox"/>	12A12	Patient 12 IAVI 57	119
<input type="checkbox"/>	15e	N70	109
<input type="checkbox"/>	17b	N70	214
<input type="checkbox"/>	2191		100
<input type="checkbox"/>	2219		100
<input type="checkbox"/>	2537		100
<input type="checkbox"/>	2558		100
<input type="checkbox"/>	2F5		546
<input type="checkbox"/>	2G12		519
<input type="checkbox"/>	3019		21
<input type="checkbox"/>	3074		100
<input type="checkbox"/>	3694		21

**Viruses**

# of Viruses = 634 (450 seqs available)

Select	Name	Subtype	# of Abs tested	Seq data
<input type="checkbox"/>	0013095_2_11	C	38	Yes
<input type="checkbox"/>	001428_2_42	C	38	Yes
<input type="checkbox"/>	0041_V3_C18	C	5	Yes
<input type="checkbox"/>	0077_V1_C16	C	24	Yes
<input type="checkbox"/>	00836_2_5	C	26	Yes
<input type="checkbox"/>	0260_V5_C1	A1	10	Yes
<input type="checkbox"/>	0260_V5_C36	A1	34	Yes
<input type="checkbox"/>	0330_V4_C3	A1	29	Yes
<input type="checkbox"/>	0439_V5_C1	A1	26	Yes
<input type="checkbox"/>	0735_V4_C1	AC	5	Yes
<input type="checkbox"/>	0815_V3_C3	ACD	36	Yes
<input type="checkbox"/>	0907_V4_C12	AD	9	Yes
<input type="checkbox"/>	0921_V2_C14	C	20	Yes
<input type="checkbox"/>	0984_V2_C2	C	5	Yes
<input type="checkbox"/>	1006_11_C3_1601	B	27	Yes
<input type="checkbox"/>	1012_11_TC21_3257	B	27	Yes

**Options**

Retrieve ☐ Antibody details ☐ Virus details ☐ Assay

or

Analyze along with virus sequences ☒ IC50 ☐ IC80 ☐ Both

Large sets of data run slowly. Limit the number of antibodies or viruses for quicker response.

☐ Exclude viruses having no sequence data

**Acknowledgements**

- This tool was inspired by the research of Dr. Anthony West ([West et al., 2013](#)) and includes data he collected from public sources.
- Tool interface was designed by Hyejin Yoon, Jennifer Macke, and Karina Yusim.
- Additional Resources
  - [LANL Neutralizing Antibody Resources](#)
  - [bNABer](#): Broadly Neutralizing Antibodies Electronic Resource
  - [SCHARP](#): Completed CAVD Studies

- Inspired by [Anthony West](#) (*West et al., PNAS 2013*) and includes data he collected from published sources
- Designed by [Hyejin Yoon](#), [Jennifer Macke](#), [Bette Korber](#), [Karina Yusim](#)

**Interface combines:**

- Antibody-virus neutralization data**
  - Env sequence data superimposed with IC50 or IC80 values
  - Antibody potency and breadth summarized over multiple studies
- Alignments and virus data**
  - Subtype, country, accession, neutralization tier, virus names
  - Patient health status, risk factor
- Antibody data**
  - Isolation study, donor ID, clonal lineage, Immuno DB records
  - PDB structure, Ab sequences
  - Neutralizing antibody features, contexts and contact residues
- Analysis per AA position**
  - AA composition, N-glycosylation sites, basic statistics
  - What is known about this position in terms of entropy, functional domain, neutralizing antibody contexts, Ab contact residues, signature predictions

### CATNAP

Compile, Analyze and Tally NAB Panels

**Purpose:** To provide easy access to data associated with HIV-1 neutralizing antibodies, including neutralization panel data, sequences, and structures. See [Explanation](#).

Select by antibody and virus
Select by study

**Studies**

# of Studies = 19

Select	Name	# of Abs tested	# of viruses tested
<input type="checkbox"/>	Andrabi et al., Virology 439:81 (2013)	18	21
<input type="checkbox"/>	Chuang et al., J Virol. 87:10047 (2013)	21	181
<input type="checkbox"/>	Diskin, et al., J. Exp. Med. 210, 1235 (2013)	6	118
<input type="checkbox"/>	Diskin, et al., Science 334, 1289 (2011)	2	82
<input type="checkbox"/>	Doria-Rose, et al., J. Virol. 86, 3393 (2012)	4	207
<input type="checkbox"/>	Falkowska et al., J Virol 86:4394 (2012)	5	244
<input type="checkbox"/>	Georgiev, et al., Science 340, 751 (2013)	4	180
<input type="checkbox"/>	Hioe et al., PLoS One 5:e10254 (2010)	7	100
<input type="checkbox"/>	Huang, et al., Nature 491, 406 (2012)	8	181
<input type="checkbox"/>	Liao et al., Nature 496:469 (2013)	4	118
<input type="checkbox"/>	Montefiori CAVD CA-VIMC	5	378
<input type="checkbox"/>	Mouquet, et al., PNAS 109, E3268 (2012)	3	119
<input type="checkbox"/>	Scheid, et al., Science 333, 1633 (2011)	6	119
<input type="checkbox"/>	Seaman, et al., J Virol 84:1439 (2010)	3	109
<input type="checkbox"/>	Walker, et al., Nature 477, 466 (2011)	20	44
<input type="checkbox"/>	Walker et al., Science 326:285 (2009)	7	159

**Options**

Retrieve ☐ Antibody details ☐ Virus details ☐ Assay

or

Analyze along with virus sequences ☒ IC50 ☐ IC80 ☐ Both

Large sets of data run slowly. Limit the number of antibodies or viruses for quicker response.

☐ Exclude viruses having no sequence data

- Inspired by [Anthony West](#) (*West et al., PNAS 2013*) and includes data he collected from published sources
- Designed by [Hyejin Yoon](#), [Jennifer Macke](#), [Karina Yusim](#)

**Interface combines:**

- Antibody-virus neutralization data**
  - Env sequence data superimposed with IC50 or IC80 values
  - Antibody potency and breadth summarized over multiple studies
- Alignments and virus data**
  - Subtype, country, accession, neutralization tier, virus names
  - Patient health status, risk factor
- Antibody data**
  - Isolation study, donor ID, clonal lineage, Immuno DB records
  - PDB structure, Ab sequences
  - Neutralizing antibody features, contexts and contact residues
- Analysis per AA position**
  - AA composition, N-glycosylation sites, basic statistics
  - What is known about this position in terms of entropy, functional domain, neutralizing antibody contexts, Ab contact residues, signature predictions



Automatically submit sequences in a batch to the HIV sequence DB search

# CATNAP: Analyze

CATNAP: IC<sub>50</sub>/HIV-1 alignment

Collapse or expand details from individual studies

2 antibodies selected to search  
Expand virus info  
More virus info in HIV Seq DB

\* Geometric means (Collapse)

● 0.1 ● <1 ● <10 ● ≥50 ○ > cutoff (μg/ml)

Virus name	10E8	Per study	PG9	Per study
0013095_2_11	0.006*	0.003, 0.013	0.013*	0.003, 0.023
001428_2_42	1.375*	0.75, 2.52	0.013*	0.003, 0.023
0077_V1_C16	1.543*	1.19, 2	0.013*	0.003, 0.023
00836_2_5	0.402*	0.492, 0.329	0.008*	0.008, 0.018
0260_V5_C36	0.598*	0.96, 0.25	2.180*	2.18, 2.18, 2.18
0330_V4_C3	1.045*	1.27, 0.86	0.018*	0.018, 0.018, 0.018
0439_V5_C1	1.185*	0.924, 1.52	>50*	>50, >50, >50
0815_V3_C3	0.343*	0.19, 0.618	>50*	>50, >50, >50
0921_V2_C14	0.909*	0.773, 1.07	-	-
1005_11_C3_1601	-	-	0.306	-
1012_11_Tc21_3257	-	-	0.372	-
1054_07_Tc4_1499	-	-	>50	-
1056_10_Ta11_1826	-	-	0.339	-
16055_2_3	0.720*	0.571, 0.909	0.013*	0.013, 0.013, 0.013
16845_2_22	0.012*	0.001, 0.14	2.380*	2.38, 2.38, 2.38
16936_2_21	0.231*	0.166, 0.321	>50*	>50, >50
211_9	-	-	0.340	-
231965_C1	0.232*	0.03, 2.23	1.510*	1.51, 1.51, 1.51
235_47	0.070*	0.077, 0.064	0.322*	0.322, 0.322, 0.322
242_14	1.207*	1.35, 1.08	0.025*	0.025, 0.025, 0.025
247_23	0.373*	0.368, 0.378	0.195*	0.195, 0.195, 0.195
Geometric mean of detected	0.273	-	0.104	-
Geometric mean of all (undetected set to 100)	0.301	-	0.695	-
% detected (detected/total)	98% (178/181)	-	79% (280/355)	-

# of viruses found: 359  
# of Abs found: 2  
# of studies found: 6  
Chuang2013 Doria-RoseNA2012 Falkowska2012 Huang2012a Walker2009a Walker2011

Antibody contact position(s) (based on HXB2)  
(See Spreadsheet of neutralizing antibody contexts and features .xls) for more information)

- 10E8 contacts: N671 W672 F673 T676 W680 K683
- PG9-like contacts: N156 N160 I165 G167 K168 V169 Q170 K171 Y173

Contact positions

Analyze HXB2 position 160 for Ab PG9

Pick Ab and click on contact position to analyze, or enter your own position

Download aa na in Fasta



## Analysis at HXB2 position 160 for Ab PG9

### Amino Acid Counts

AA	Count	# for detected	# for undetected	Fisher test p-value
N	208	175	33	1.187e-09
X	3	2	1	0.5149
K	3	0	3	0.009233
D	3	0	3	0.009233
S	2	0	2	0.04476
Y	2	0	2	0.04476
R	2	0	2	0.04476
H	1	0	1	0.2133
-	1	0	1	0.2133
Total	225	177	48	
no seq	130			
Grand total	355			

### N-glycosylation Motif Counts

NxST	Count	# for detected	# for undetected	Fisher test p-value
g+	204	174	30	3.671e-11
g-	20	3	17	2.086e-10
-	1		1	
Total	225	177	48	
no seq	130			
Grand total	355			

### HXB2

IEKGEIKNCSFNIISTIRG-KVQKEYAFFYKLDIIPIDN-----DTT  
|-----|-----|-----|-----|-----|-----|-----|-----|  
0-----160-----170-----180-----190-----200

YKEDIRNCSFNATTEVKD-KKQKVHALFYRLDIVPLNKRNSSESEEN-----SSG  
'NGDEMKNCSFNIITTEIRD-KKQKAYALFYRLDIVPLERENRGDSN-----SAS  
'TSNEMKNCSFNIITTEIRD-KKKESAIKYKLDVPLDNGNSG-----NYS  
'TYESMKNCSFNIITTELD-KKQSVYALFYRLDIVPLNN-----SNE  
'MEGEIKDCSFNIITTELD-KRQKVHSLFYRLDIVQINSQT-----NSS  
'TRDELRCNCSFNIITTELD-RRQKVHSLFYRLDIVEIENRTNRT-----NNT  
'TENERKNCSFNIITTELD-KSKQVYSLFYRLDIVDSDNSD-----NSN  
'STADMKNCSFNVPTAIRD-RKQKVYSLFYRLDIVQIDKKNDNSNS-----NIT  
'---IMTNTCFNIITTELD-KKRKASAFYRLDIVPLNGDSNGS-----SSG  
'DKGEMKNCSFNIITTSIRG-KMQKEYALFYKLDIVPIDNGKND-----TNT  
'ESGEIKNCSFNIITTSVRD-KVQKEYALFYKLDIVPITN-----ESS  
'DPGEIKNCSFNIATPIKD-KRQKEYALFYKSDVVPIDEDN-----DTT  
'IEKGEIKNCSFNIITNIRD-KYQKAYALFYKLDVVPIDEDNATGN-----DTF  
'NGEIKNCSFNIATTEIRD-KKQKVYALFYRLDIVPLEERKG-----NSS  
'DMGEIKNCSFNIITTELD-KKQKVHALFYRLDIVSLEKDNSSKND-----SNE  
'INVEEMKNCSFNIITTELD-RKQTVYALFYRLDIVPLNENKST-----SSE  
'MEGEIKNCSFNIITTELD-KNQKVYALFYRQDVQIGNNN-----NSS  
'PEAGMKNCSFNIITTEVKD-KKKLVYAHFYRLDIVQLDGD-----NTN  
'QGEEMKNCSFNIITAEIRD-KRKNEYALFYRQDVQINET-----DNS  
'IMKGEITNCSFNIITTELD-KKQKVSAFFYRQDVVPVNSNQ-----DNS  
'INTEDMKNCSFNIITIVRD-KKKQYALFYRLDIVEINP-----NDT  
'CKNKNCSFNIITTELD-KKQKVYALFYRLDIVPLEERKG-----SSG

### About this position

Position: Env 160 (193 in alignment above)

Entropy, M group: 0.401

Functional domain: gp120 (Kwong2000), V2 (Leonard1990)

### Antibody features of this position

Mutation affects PG9-like Ab sensitivity: Loss of glycan confers resistance; PG9-like class includes PG16, PGT141, 145, CH01-CH04 (V1V2 glycan, Doria-RoseNA2012)

PG16 signature predictions: PG16: glycosylation at N160 is associated with increased susceptibility to neutralization; intermediate quality of support. (V1V2 glycan, West2013)

PG9-like contacts: PG9 glycan contact; PG9-like class includes PG16, PGT141, 145, CH01-CH04 (V1V2 glycan, McLellan2011)

PG9 signature predictions: PG9: N160 is associated with increased susceptibility to neutralization; intermediate quality of support. (V1V2 glycan, West2013)

(For more information, check Neutralizing Antibody Contexts & Features)

Position highlighted

## Antibody information

Number of antibodies: 2

Download ☒ antibody aa sequences ☐ antibody na sequences in 

Download table below

## CATNAP: Ab info

Antibody	Structure	Donor	Clonal lineage	Isolation paper	Neutralizing antibody feature	Heavy chain	Light chain
<a href="#">PGT125</a>		Donor 36	PGT128	<a href="#">Walker2011</a>	<input type="radio"/> <a href="#">PGT121-123, 125-128, 130 and 131</a>	<a href="#">JN201897</a> Synthetic construct isolate PGT125 anti-HIV immunoglobulin heavy chain variable region gene, partial cds QSQLQESGPRLVEASETLSTCNVSGESTGACTYFHWGVRQAPKGLWIGSLBRCQSFWSGWTFFNPSLKSRLTISLDTFKNQVFLKLTSLTAADTATYTCARFGGEVLVYNRWPKPAWVDLWGRGIPVTVTS	<a href="#">JN201914</a> Homo sapiens isolate PGT125 anti-HIV immunoglobulin light chain variable region mRNA, partial cds QSALTTPPSASGSPGQSITISCTGATNFVSWYQQFPDKAPKLIIFGVDRPQGVDFRFGSGRSGTTASLTSLRQLTDDAEVYCYCSLVGNWDVIFGGGTITVL
<a href="#">PGT127</a>	<input type="radio"/> <a href="#">PGT127 complexed with Man(9)</a>	Donor 36	PGT128	<a href="#">Walker2011</a>	<input type="radio"/> <a href="#">PGT121-123, 125-128, 130 and 131</a> <input type="radio"/> <a href="#">PGT127 / PGT128</a>	<a href="#">JN201899</a> Homo sapiens isolate PGT127 anti-HIV immunoglobulin heavy chain variable region mRNA, partial cds QSQLQESGPRLVEASETLSTCNVSGESTGACTYFHWGVRQAPKGLWIGSLBRCQSFWSGWTFFNPSLKSRLTISLDTFKNQVFLKLTSLTAADTATYTCARFGGEVLVYNRWPKPAWVDLWGRGIPVTVTS	<a href="#">JN201916</a> Homo sapiens isolate PGT127 anti-HIV immunoglobulin light chain variable region mRNA, partial cds QSALTTPPSASGSPGQSITISCTGATNFVSWYQQFPDKAPKLIIFGVDRPQGVDFRFGSGRSGTTASLTSLRQLTDDAEVYCYCSLVGNWDVIFGGGTITVL

Link to structure in PDB

Description PGT121-123, 125-128, 130 and 131

Antibody class V3 glycan

Reference [Walker2011](#)

Type antibody

MAb name [PGT121](#) [PGT122](#) [PGT123](#) [PGT125](#) [PGT126](#) [PGT127](#) [PGT128](#) [PGT130](#) [PGT131](#) (Click MAb name to get to Immunology DB notes)

Env pos.	Feature	AA	Entropy	Annotation
301	gp120, V3 loop	N	0.261	N-linked glycans in positions 332 and/or 301 were important for neutralization by PGT MAbs 125-128, 130, and 131
332	gp120	N	0.904	N-linked glycans in positions 332 and/or 301 were important for neutralization by PGT MAbs 125-128, 130, and 131, this glycan is also critical for PGT 121-123.

Important position(s) with Hxb2 amino acid: N301 N332

Alamos  
LABORATORY

## Virus information

Number of viruses: 14

Download table below

Automatically submit all selected sequences in a batch to the HIV sequence search interface

## CATNAP: Virus info

More info in HIV Sequence DB

Virus name	Subtype	Country	Patient health	Days post infection	Days from seroconversion	Fiebig	Risk factor	Accession	Tier	Alias	HIV DB name	Seq data	LANL comments
216_F2_E3_5	A1C	TANZANIA				6	Heterosexual	<a href="#">HM215277</a>			216_F2_E3_5	Yes	
231965_C1	D	UGANDA	acute infection		early	1 or 2		<a href="#">JQ361079</a>	2	231965, 231965_C01	231965_c01	Yes	
231966_C2	D	UGANDA	acute infection		early	1 or 2		<a href="#">JX512899</a>	2	231966_C02	231966_c02	Yes	
234_F1_16_57	C	TANZANIA			early	5	Heterosexual	<a href="#">HM215278</a>			234_F1_16_57	Yes	
235_47	02_AG	CAMEROON				6	Not Recorded	<a href="#">EU513195</a>	2	235	235	Yes	Sequence does not match accession. This sequence/clone was the one used in neutralization studies but it has not yet been deposited in GenBank.
242_14	02A1	CAMEROON				6		<a href="#">EU513188</a>	1B or 2	242	242	Yes	
246_F3_C10_2	AC	TANZANIA				6	Heterosexual	<a href="#">HM215279</a>			246_F3_C10_2	Yes	
246F_C1G	C	ZAMBIA	acute infection		early	2	Heterosexual	<a href="#">FJ496194</a>	2	ZM246, 246F	ZM246F_f1D5	Yes	
247_23	DU	CAMEROON					Not Recorded	<a href="#">EU683891</a>	2	247	247	Yes	
249M_B10	C	ZAMBIA	acute infection		early		Heterosexual	<a href="#">EU166862</a>	2	249M	ZM249M_080503_SGA_B10	Yes	
25710_2_43	C	INDIA	acute infection		45	5	Heterosexual	<a href="#">EF117271</a>	1B or 2	25710	HIV_25710_2	Yes	
25711_2_4	C	INDIA	acute infection		45	3	Heterosexual	<a href="#">EF117272</a>	1B or 2	25711	HIV_25711_2	Yes	
25925_2_22	C	INDIA	acute infection		45	3	Heterosexual	<a href="#">EF117273</a>	1B or 2	25925	HIV_25925_2	Yes	
26191_2_48	C	INDIA	acute infection		45	3	Heterosexual	<a href="#">EF117274</a>	2	26191	HIV_26191_2	Yes	

Link to the sequence record in the HIV Sequence DB

# CATNAP: Assay info

## Assay

Number of antibodies selected: 82 (10-1074, 10-996, 10E8, 12A12, 15e ...)

Number of viruses selected: 14 (216\_F2\_E3\_5, 231965\_C1, 231966\_C2, 234\_F1\_16\_57, 235\_47 ...)

Number of studies selected: 0

Number of data found: 349

[Download](#) table below with additional virus info

### Hide virus info

Antibody	Virus	Subtype	Tier	Country	Accession	Alias	Reference	IC50	Mean IC50	IC80	Mean IC80
10-1074	231965_C1	D	2	UGANDA	JQ361079	231965_C01, 231965	<a href="#">Mouquet, et al., PNAS 109, E3268 (2012)</a>	>50	>50		
10-1074	231966_C2	D	2	UGANDA	JX512899	231966_C02	<a href="#">Mouquet, et al., PNAS 109, E3268 (2012)</a>	>50	>50		
10-1074	235_47	O2_AG	2	CAMEROON	EU513195	235	<a href="#">Mouquet, et al., PNAS 109, E3268 (2012)</a>	0.05	0.050		
10-1074	246F_C1G	C	2	ZAMBIA	FJ496194	246F, ZM246	<a href="#">Mouquet, et al., PNAS 109, E3268 (2012)</a>	0.022	0.022		
10-1074	249M_B10	C	2	ZAMBIA	EU166862	249M	<a href="#">Mouquet, et al., PNAS 109, E3268 (2012)</a>	>50	>50		
10-996	231965_C1	D	2	UGANDA	JQ361079	231965_C01, 231965	<a href="#">Mouquet, et al., PNAS 109, E3268 (2012)</a>	>50	>50		
10-996	231966_C2	D	2	UGANDA	JX512899	231966_C02	<a href="#">Mouquet, et al., PNAS 109, E3268 (2012)</a>	>50	>50		
10-996	235_47	O2_AG	2	CAMEROON	EU513195	235	<a href="#">Mouquet, et al., PNAS 109, E3268 (2012)</a>	0.128	0.128		
10-996	246F_C1G	C	2	ZAMBIA	FJ496194	246F, ZM246	<a href="#">Mouquet, et al., PNAS 109, E3268 (2012)</a>	0.092	0.092		
10-996	249M_B10	C	2	ZAMBIA	EU166862	249M	<a href="#">Mouquet, et al., PNAS 109, E3268 (2012)</a>	>50	>50		
10E8	231965_C1	D	2	UGANDA	JQ361079	231965_C01, 231965	<a href="#">Chuang et al., J Virol. 87:10047 (2013)</a> <a href="#">Huang, et al., Nature 491, 406 (2012)</a>	2.23 8.03	4.232		
10E8	235_47	O2_AG	2	CAMEROON	EU513195	235	<a href="#">Chuang et al., J Virol. 87:10047 (2013)</a> <a href="#">Huang, et al., Nature 491, 406 (2012)</a>	0.064 0.077	0.070		
10E8	242_14	O2A1	1B or 2	CAMEROON	EU513188	242	<a href="#">Chuang et al., J Virol. 87:10047 (2013)</a> <a href="#">Huang, et al., Nature 491, 406 (2012)</a>	1.08 1.35	1.207		

 **Los Alamos**  
NATIONAL LABORATORY

## Neutralization panel: looking forward

- Neutralization data from many more studies, particularly IC80 values.
- Additional measurements of neutralization and antibody binding.
- Alignments of antibody variable domain sequences.
- Signature analysis results using collected data.
- Autologous neutralization data for studies with multiple HIV sequences and multiple antibodies isolated from the same donor.

 **Los Alamos**  
NATIONAL LABORATORY



# ELF

- If you have a peptide that reacts with CD8+ T cells from a person with known HLA type, enter:
  - ☐ The peptide that reacts with CD8+ T-cells
  - ☐ The HLA type of the person with the reactive CD8+ T cells
- ELF will help identify the possibly reactive epitope by
  - ☐ Highlighting appropriate HLA anchor motifs in the peptide
  - ☐ Aligning all known epitopes embedded in the peptide from the database to your query sequence, with links to epitope entries
  - ☐ Finding potential epitopes based on Immune Epitope Database (IEDB) binding predictions <http://www.immuneepitope.org/>
- Other useful information provided:
  - ☐ Genomic location of your peptide
  - ☐ Database records for known CTL epitopes in this region, regardless of HLA.



## ELF

### Epitope Location Finder

**Purpose:** search a submitted protein sequence for (1) known epitopes from our immunology databases, (2) epitopes predicted by consensus binding motifs, and (3) epitopes predicted by the IEDB binding algorithm. For details see [ELF Explanation](#).

**Input**

Paste [protein sequence](#)  <50 amino acids, raw format

**Options**

Show [known epitopes](#) ☒ from CTL and Helper databases

Find potential epitopes ☒ based on [anchor residues](#)

Choose HLA(s) (Class I and Class II)

Use control-click for multiple selection

By genotype

- A\*3004
- A\*3101
- A\*3201
- A\*3303
- A\*6601
- A\*6801
- A\*6802

By serotype

- A33(19)
- A69(28)
- A68(28)
- A30(19)
- A66(10)
- A2

HLA selection is synchronized between 2 analysis options

Find potential epitopes ☒ based on [IEDB binding predictions](#)

Choose HLA(s) or MHC(s) (synchronized with genotype selections above)

HLA Class I

- A\*6611
- A\*6612
- A\*6613
- A\*6614
- A\*6615
- A\*6801
- A\*6802

HLA Class II

- DRB3\*0221
- DRB3\*0225
- DRB3\*0301
- DRB3\*0303
- DRB4\*0101
- DRB4\*0103
- DRB5\*0101

Animal MHC Class I

chimpanzee

- Patr-A\*0101
- Patr-A\*0201
- Patr-A\*0301
- Patr-A\*0302
- Patr-A\*0401
- Patr-A\*0402

Animal MHC Class II

mouse

- H2-IAb
- H2-IAd
- H2-IEd

Display binders ☒ Show 1 best binder(s) per MHC


☐ Show below 20 percentile rank (1-100) per MHC

E-mail result ☐ Predictions are slow. For more than a few HLAs/MHCs, we recommend e-mailed result.

You can choose how many top binders to show per MHC, or use a binding percentile rank cutoff




## ELF results 1:

### Epitopes from our CTL database aligned to your query sequence

Bold red letters indicate residues that differ from the query sequence. The symbol  means the HLA of the epitope matches one of your submitted HLAs. Click on the epitope to see full database entry. Click on "align" to align the epitope to the sequence database via QuickAlign.

Epitopes shown here are completely within the bounds of your query. Epitopes that overlap the ends of your query are included in the "View database records" links above.

Download this alignment in format

```
DTVLEDMNLPGRWKPKMIG
DTVLEEMNL A*6802 align 
DTVLEDINL A*6802 align 
DTVLEEWNL A*6802 align 
DTVLEEMNL A68 align
DTVLEEMNL A28 align
DTVLEDMNL align
EEMNLPGRW B44 align
EINLPGRW B44 align
EEMNLPGRW B*4402 align
EEMNLPGRW B*4403 align
EEMNLPGRW B18,B40,B44 align
EDMNLPGRW align
EEMNLPGRW B*44 align
EINLPGRW B*4403 align
EEMNLPGRW align
LPGRWKPKMI Cw3 align
LPGRWKPKMI B7 align
```

Clicking on an epitope takes you to respective CTL or Helper epitope Database entries

## ELF results 2:

### Potential epitopes based on anchor residues

These peptides have C-terminal anchor residues, highlighted in blue, and internal anchors highlighted in magenta. These anchor residues match one or more motifs associated with the submitted HLA.

Download this alignment in format

```
DTVLEDMNLPGRWKPKMIG
DTVLEDMNL (A*0205 .....[L])
DTVLEDMNL (A*6802 .[TV].....[VL])
TVLEDMNLP (A*0206 .[VQ].....)
LEDNLPGR (DRB5*0101,DRB5*0101 [FYLM]..[QVIM]....[RK])
```

# ELF results 3:

## Potential epitopes based on IEDB database MHC binding predictions, by Alexander Sette's group

### Potential epitopes based on IEDB binding predictions

Top binders for each MHC are highlighted in blue.

Prediction method: IEDB recommended

Low percentile = good binders

Show up to 1 binder(s) per MHC

#### Class I

Selected allele(s): A\*6802, B\*1501

Download this alignment in format

DTVLEDMNLPGRWKPKMIG (Click MHC to see full list of IEDB predictions for that MHC)

DMNLPGRW B\*1501 (26)

MNLPGRWK A\*6802 (3.0)

Clicking on MHC links to the full list of IEDB predictions for that MHC (see next slide)

#### Class II

Selected allele(s): DRB5\*0101

Download this alignment in format

DTVLEDMNLPGRWKPKMIG (Click MHC to see full list of IEDB predictions for that MHC)

TVLEDMNLPGRWKPK DRB5\*0101 (17.17)



# ELF results 3:

## Potential epitopes based on IEDB database MHC binding predictions, by Alexander Sette's group

### IEDB Analysis Resource

Home Help Example Reference Download Contact

### MHC-I binding predictions - Prediction Results

#### Input Sequences

#	Name	Sequence
1	sequence 1	DTVLEDMNLPGRWKPKMIG

Prediction method: IEDB recommended | Low percentile = good binders

Check to expanded the result: ☐

Allele	#	Start	End	Peptide Length	Sequence	Method used	Percentile Rank
HLA-B*15:01	1	6	13	8	DMNLPGRW	NetMHCpan	26
HLA-B*15:01	1	3	13	11	VLEDMNLPGRW	NetMHCpan	27
HLA-B*15:01	1	3	11	9	VLEDMNLPG	Consensus (ANN,SMM,CombLib_Sidney2008)	27.60
HLA-B*15:01	1	8	17	10	NLPGRWKPKM	NetMHCpan	31
HLA-B*15:01	1	7	17	11	MNLPGRWKPKM	NetMHCpan	35
HLA-B*15:01	1	2	9	8	TVLEDMNL	NetMHCpan	36
HLA-B*15:01	1	2	11	10	TVLEDMNLPG	NetMHCpan	47
HLA-B*15:01	1	4	11	8	LEDNMNLPG	NetMHCpan	48





# Mosaic vaccine tools

**Mosaic Vaccine Designer:** The Mosaic Vaccine Designer will generate candidate vaccine protein 'cocktails' that optimize coverage of potential T-cell epitopes found in a given background set of protein sequences.

**Epitope Coverage Assessment:** Alignment independent “n-mer” coverage of sequences by vaccines or peptides.

**Positional Epitope Coverage Assessment:** Alignment dependent coverage of sequences by vaccines or peptides.



## Mosaic Vaccine Designer

**HIV sequence database**

DATABASES SEARCH ALIGNMENTS TOOLS PUBLICATIONS GUIDES Search Site

### Mosaic Vaccine Designer

Purpose: The Mosaic Vaccine Designer will generate candidate vaccine protein cocktails that optimize the coverage, by a small set of mosaic proteins that could be included in a vaccine cocktail, of potential T-cell epitopes in a large diverse set of proteins. The resulting mosaic proteins in the proposed vaccine cocktail resemble viral proteins from the input set of natural viral proteins (the training set), but are assembled from fragments of the natural proteins using a genetic algorithm (a computational optimization method). This method was first applied to HIV, but is readily generalized and could be applied to other variable pathogens.

Functions:

- Create mosaic sequence cocktail runs the genetic algorithm to generate a cocktail of synthetic sequences with near-optimal coverage
- Pick the best natural sequences selects unmodified natural sequences from the training set in order of coverage
- See the coverage distribution of natural sequences shows the coverage of a randomly selected set of natural sequence cocktails

Usage: Paste your protein sequences in the box below, or upload a file containing sequences. Most common [sequence formats](#) are accepted. As soon as your job is completed, a link to your results will be sent to your email address which you provided. To manage more detailed parameters, go to the Advanced Input. (Your job may take several hours or even days, according to your input.)

Related Programs:

- [Epitope Coverage Assessment Tool-EpiCover](#)
- [Positional Epitope Coverage Assessment Tool-PosCover](#)

Reference: [Divalent vaccine design article](#) | [PubMed version](#)

**Input**

A1.CH...  
MGDNHSRSSLVQVFIERKHAPPTPTPTPAAGVGAVSQDLAKHGA1  
A1.XE.99a  
MGDNHSRSSLVQVFIERKHAPPTPTPTPAAGVGAVSQDLAKHGA1  
A1.XE.99b  
MGDNHSRSSLVQVFIERKHAPPTPTPTPAAGVGAVSQDLAKHGA1

Paste set of protein sequences

☒ Sample Input

Or upload protein sequence file

**Options** Basic Advanced

Function ☒ Create mosaic sequence cocktail  
☐ Pick the best natural sequences  
☐ See the coverage distribution of natural sequences

Cocktail Size (1-10)

Epitope Length (8-12)

Rare Threshold

Paste fixed sequences

Or upload fixed sequence file

Last modified: Wed Jun 9 12:50 2009

Input: protein sequence set for a target population, does not need to be aligned.

Number of mosaic proteins in the set.

Epitope length.



# Epitope Coverage Assessment - Epicover

## Input

Use output from MakeVaccine tool

Provide a job number to access output from the [Mosaic Vaccine Designer](#) tool:

OR

Provide input sequences

Paste antigen protein  
sequence(s):  
[\[Sample Input\]](#)

upload more [ + ] antigen sequence files

and/or upload as files:

[Browse...](#)

Paste test set protein  
sequences:

upload more [ + ] test sequence files

and/or upload as files:

[Browse...](#)

## Options

Send results as an email instead of displaying in browser  
(useful in case of a browser time-out): ☐

Nominal epitope length:

Maximum amino acid mismatches to score (range from 0):

Minimum number of occurrences of a potential epitope  
in viral protein set to consider for coverage:

Precision to use when reporting coverage:  decimal places

## Advanced Options

Upload file of grouped sequence names

[Browse...](#)

Report on subsets defined according to first  character(s) in sequence names

[Submit](#) [Reset](#)

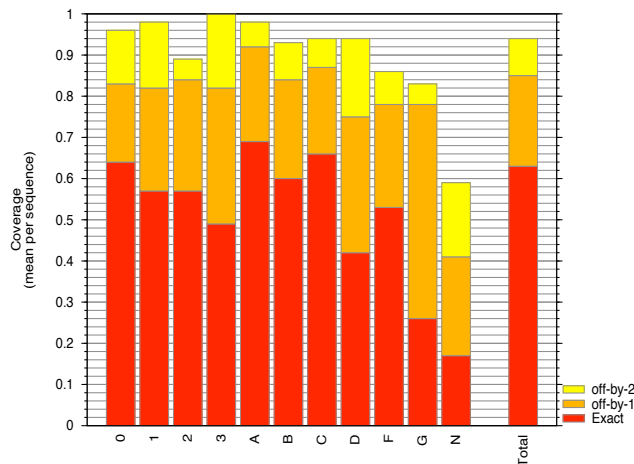
**Input:**  
Vaccine set  
Test set

Can report on  
subsets defined  
according to the  
first several  
characters in  
sequence  
names or  
user-defined  
subsets



## Epicover output

vaccine set	subset	subset count	Off-by-0	Off-by-1	Off-by-2	rare(<3,>1)	unique	absent
vaccine_set_from_user	Total	63	0.6615	0.8914	0.9660	104	114	334
vaccine_set_from_user	A	11	0.7232	0.9429	0.9935	47	36	334
vaccine_set_from_user	B	11	0.6378	0.8845	0.9755	25	19	334
vaccine_set_from_user	C	35	0.6921	0.8994	0.9637	51	45	334
vaccine_set_from_user	D	4	0.4217	0.7546	0.9443	4	9	334
vaccine_set_from_user	F	1	0.5300	0.7800	0.8600	4	5	334
vaccine_set_from_user	G	1	0.2597	0.7792	0.8312	0	0	334



# Positional Epitope Coverage Assessment - Posicover

Provide a job # from   
[Mosaic Vaccine Designer](#): (Only the antigen set is used. Provide the ALIGNED viral test set below)  
 AND/OR  
 Paste antigen protein set or peptide cocktail: (alignment not required)  
[\[Sample Input\]](#)  
  
 upload more [ + ] antigen files  
 and/or upload antigen file(s):

**Test set proteins**

Paste ALIGNED test viral protein set:  
[\[Sample Input\]](#)  
  
 or upload an ALIGNED test proteins file:

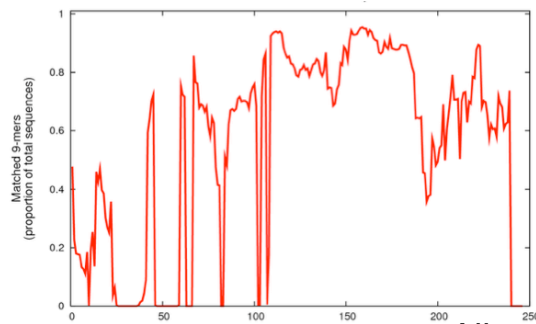
Input:

Vaccine set  
 ALIGNED test set

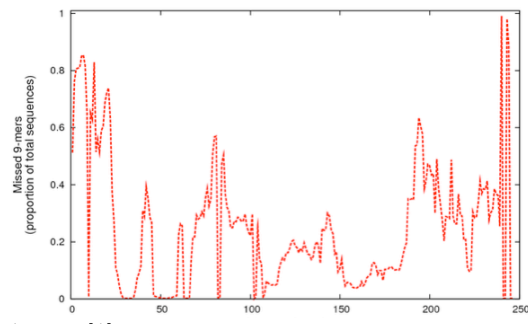


## Examples of Posicover outputs

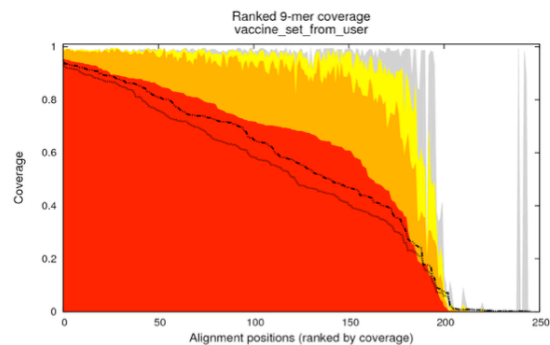
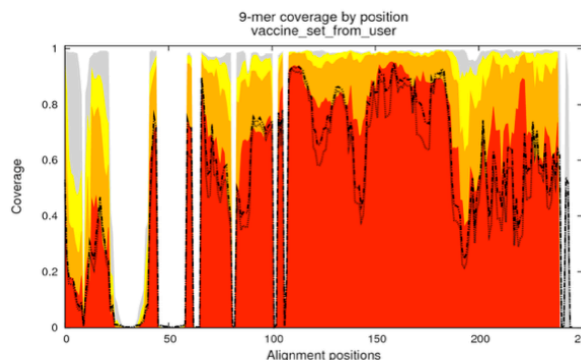
Matched 9-mers



Missed 9-mers



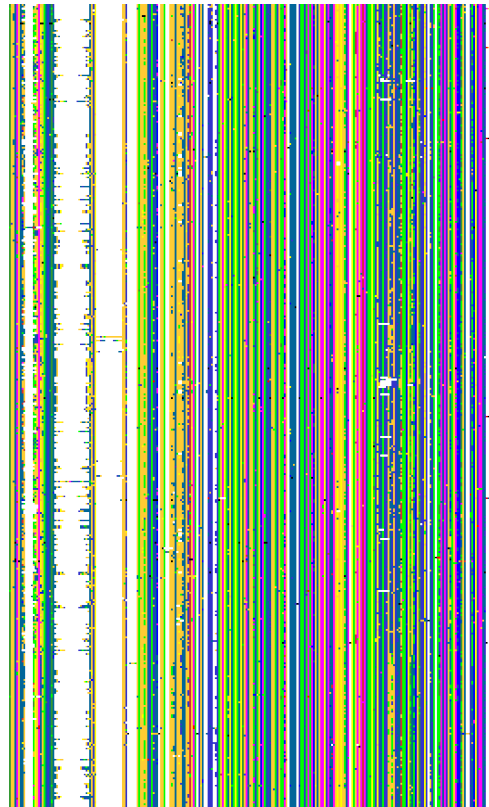
Alignment positions



## Examples of Posicover outputs

User's sequence alignment:

Each aa is represented as a single-colored square



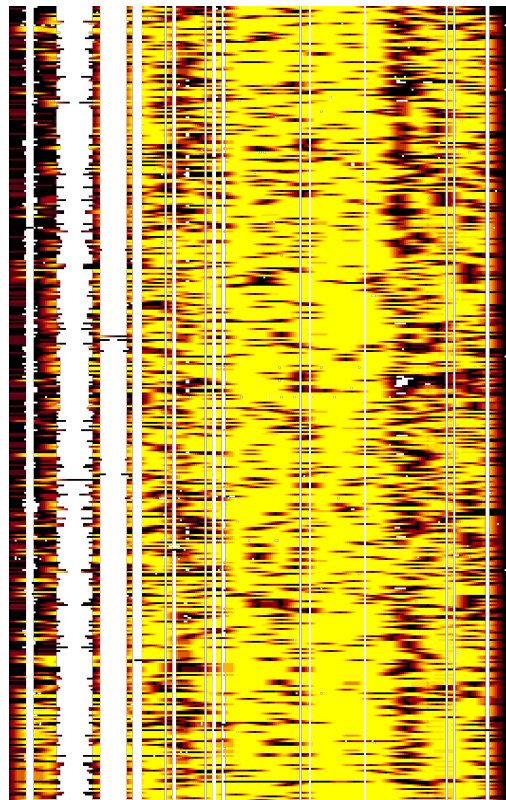
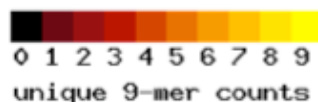
## Examples of Posicover outputs

Each amino acid is colored according to the set of 9-mers that contain it:

Yellow: all 9-mers that overlap with amino acid are perfectly matched in a test vaccine set;

Increasingly red: fewer and fewer matches in the overlapping set of 9-mers that span the amino acid;

Black: amino-acid residues that are not included in any vaccine set



**Please leave any comments or  
suggestions with us:**

